

# **BLASTZ-WEB: a web interface to the BLAST-Z algorithm**

ID - 230

Traini Alessandra<sup>1</sup>, Chiusano Maria Luisa<sup>2</sup>

<sup>1</sup>traini.alessandra@gmail.com

<sup>2</sup>Department of Soil, Plant, Environmental and Animal Production Sciences, University Federico II, Naples

## **Motivation**

BlastZ [1] is a pairwise alignment tool, based on the gapped BLAST algorithm. The BlastZ algorithm permits the alignment of genomic sequences and it is useful for the detection of local similarities on long nucleotide sequences as, for example, putatively functional conserved DNA sequences or syntenic regions. The BlastZ algorithm searches for short near-perfect matches, with a fixed length, between the two sequences. Each matching region is extended in two steps: first a gap-free dynamic programming approach is used, then, for matches above a specific threshold, the algorithm extends consecutive local similarities using a dynamic programming approach including gaps. Only the matches with a score above a specified threshold are reported.

The BlastZ text output file, which includes only the coordinates of the local similarities and of the possible extended alignments, is not easy to derive the sequence alignments and is rather difficult to read for non-expert users.

Commonly, BlastZ output is graphically supported by LAJ (Local Alignments with Java) [2], an interactive viewer available from the PipMaker package [3]. The tool can display an interactive dot-plot, a percent identity plot (pip), a text representations of the aligned sequences and a diagram showing the genome sequence organization, when known. However, LAJ is limited to sequences shorter than 170 Mb, and the dot-plot viewer, though useful for a general view on regions of similarity between two sequences, is not user friendly and flexible for specific comparative analysis and management of the multiple local similarities.

Therefore, in order to allow a user-friendly view of the information content resulting, we wrote a script and designed a web based interface to parse the BlastZ output file.

## **Methods**

### **Results**

The BlastZ output parser is written using Perl, while the web interface is created using HTML and PHP scripts. Input data files must be in FASTA format. BlastZ analysis is performed on two sequences at time. The pairwise sequence alignments are defined and reported in a HTML page. All the local alignments, resulted from the BlastZ gap-free step, are provided together with a popup links to a global view, where extended local alignments are joined using gaps.

Two multi-FASTA sequence files can be uploaded in order to align each sequence in the first set (queries) versus all the sequences in the second set (subjects). In this case, the results are reported as a summary list. The alignments produced between each pair of submitted sequences are all reported. For each pairwise comparison, links to the BlastZ text output as well as to the HTML viewer are included. The user can also view the same results through the LAJ graphical user interface.

References: [1] Schwartz S. et al.: Human-mouse alignments with BLASTZ. *Genome Res.* 2003, 13:103-7.

[2] LAJ was written by Cathy Riemer from Miller lab at the Penn State University Center for Comparative Genomics and Bioinformatics.

[3] Schwartz S. et al.: PipMaker: A Web Server for Aligning Two Genomic DNA Sequences. *Genome Res.* 2000, Vol. 10, 4:577-586.

**Availability:** The BlastZ web application can be freely accessed at <http://biosrv.cab.unina.it/blastzweb/>

**Email:** traini.alessandra@gmail.com